



(12) **EUROPEAN PATENT APPLICATION**
published in accordance with Art. 153(4) EPC

(43) Date of publication:
21.03.2012 Bulletin 2012/12

(21) Application number: **10774566.3**

(22) Date of filing: **14.05.2010**

(51) Int Cl.:
G10L 19/02 (2006.01) G10L 19/00 (2006.01)
H04S 3/00 (2006.01) H04H 20/88 (2008.01)
H04S 1/00 (2006.01) G10L 19/14 (2006.01)

(86) International application number:
PCT/CN2010/072781

(87) International publication number:
WO 2010/130225 (18.11.2010 Gazette 2010/46)

(84) Designated Contracting States:
**AL AT BE BG CH CY CZ DE DK EE ES FI FR GB
GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO
PL PT RO SE SI SK SM TR**

(30) Priority: **14.05.2009 CN 200910137565**

(71) Applicant: **Huawei Technologies Co., Ltd.**
Shenzhen, Guangdong 518129 (CN)

(72) Inventors:
• **ZHANG, Qi**
Shenzhen
Guangdong 518129 (CN)
• **ZHANG, Libin**
Shenzhen
Guangdong 518129 (CN)

(74) Representative: **Barth, Stephan Manuel**
Isarpatent
Postfach 44 01 51
80750 München (DE)

(54) **AUDIO DECODING METHOD AND AUDIO DECODER**

(57) Embodiments of the present invention disclose an audio decoding method, including: determining that bitstreams to be decoded are monophony coding layer and first stereo enhancement layer bitstreams; decoding the monophony coding layer to obtain a monophony decoded frequency-domain signal; reconstructing left and right channel frequency-domain signals in a first sub-band region by utilizing the monophony decoded frequency-domain signal after an energy adjustment; and reconstructing left and right channel frequency-domain signals in a second sub-band region by utilizing the monophony decoded frequency-domain signal without the energy adjustment.

Extract a channel parameter ITD, perform a channel delay adjustment on left and right channel signals according to the ITD parameter, and perform down-mixing on the adjusted left and right channel signals to obtain an M signal and an S signal	S11
↓	
Divide the frequency-domain signals of the left and right channels into 8 sub-bands, extract, according to the sub-bands, left and right channel parameters ILDs, and quantize and code the parameters to obtain the quantized channel parameters ILDs	S12
↓	
Code the M signal and perform local decoding to obtain a locally decoded frequency-domain signal M ₁	S13
↓	
Divide the M ₁ frequency-domain signal into 8 sub-bands same as those of the left and right channels, compute an energy compensation parameter of sub-bands 5, 6, and 7, and quantize and code the energy compensation parameter to obtain the quantized energy compensation parameter	S14
↓	
Perform a frequency spectrum peak value analysis on the locally decoded frequency-domain signal M ₁ to obtain a frequency spectrum analysis result	S15
↓	
Select an optimum energy adjusting factor multiplier, perform an energy adjustment on the decoded frequency-domain signal M ₁ according to formula (3) to obtain a frequency-domain signal M ₂ after the energy adjustment, and quantize and code the energy adjusting factor multiplier	S16
↓	
Compute left and right channel residual signals according to formula (4) by utilizing the frequency-domain signal M ₂ after the energy adjustment, left and right channel frequency-domain signals L and R, and the quantized channel parameters ILDs W _q of the left and right sound channels	S17
↓	
Perform a K-L transform on the left and right channel residuals, quantize and code a transform kernel H, and perform hierarchical and multiple quantizing and coding on a residual primary component and a residual subordinate component that are obtained through the transform	S18
↓	
Perform, according to the importance, hierarchical bitstream encapsulation on various coding information extracted at the coding end, and transmit the coding bitstream	S19

FIG. 1

Description

[0001] This application claims priority to Chinese Patent Application No. 200910137565.3, filed with the Chinese Patent Office on May 14, 2009, and entitled "AUDIO DECODING METHOD AND AUDIO DECODER", which is incorporated herein by reference in its entirety.

FIELD OF THE INVENTION

[0002] The present invention relates to the field of multi-channel audio coding and decoding technologies, and in particular, to an audio decoding method and an audio decoder.

BACKGROUND OF THE INVENTION

[0003] Currently, multi-channel audio signals are widely used in various scenarios, such as telephone conference and game. Therefore, coding and decoding of multi-channel audio signals is drawing more and more attention. Conventional waveform-coding-based coders, such as Moving Pictures Experts Group II (MPEG-II), Moving Picture Experts Group Audio Layer III (MP3), and Advanced Audio Coding (AAC), code each channel independently when coding a multi-channel signal. Although this method can well restore the multi-channel signal, a required bandwidth and coding rate are several times as high as those required by a monophonic signal.

[0004] Currently, popular stereo or multi-channel coding technology is parametric stereo coding, which may use little bandwidth to reconstruct a multi-channel signal whose auditory experience is completely the same as that of an original signal. The basic method is: at a coding end, down-mixing the multi-channel signal to form a monophonic signal, coding the monophonic signal independently, extracting channel parameters between channels simultaneously, and coding these parameters; at a decoding end, first decoding the down-mixed monophonic signal, and then decoding the channel parameters between the channels, and finally using the channel parameters and the down-mixed monophonic signal together to form each multi-channel signal. Typical parametric stereo coding technologies, such as the PS (Parametric Stereo), are widely used.

[0005] In parametric stereo coding, the channel parameters that are usually used to describe interrelationships between channels are as follows: Inter-channel Time Difference (ITD), Inter-channel Level Difference (ILD), and Inter-Channel Coherence (ICC). These parameters may indicate stereo acoustic image information, such as a sound source direction and location. By coding and transmitting these parameters and the down-mixed signal that is obtained from the multi-channel signal at the coding end, the stereo signal may be well reconstructed at the decoding end with a small occupied bandwidth and a low coding rate.

[0006] However, during the process of researching and implementing the prior art, the inventor of the present invention finds that: By using the conventional parametric stereo coding and decoding method, a problem that processed signals at the coding end and the decoding end are inconsistent exists, and the inconsistency of the coding and decoding signals may cause quality of a signal obtained through decoding to decline.

SUMMARY OF THE INVENTION

[0007] Embodiments of the present invention provide an audio decoding method and an audio decoder, which can enable processed signals at a coding end and a decoding end to be consistent, and improve quality of a decoded stereo signal.

[0008] The embodiments of the present invention include the following technical solutions:

An audio decoding method, including:

determining that bitstreams to be decoded are monophony coding layer and first stereo enhancement layer bitstreams;

decoding the monophony coding layer bitstream to obtain a monophony decoded frequency-domain signal; reconstructing left and right channel frequency-domain signals in a first sub-band region by utilizing the monophony decoded frequency-domain signal after an energy adjustment; and reconstructing left and right channel frequency-domain signals in a second sub-band region by utilizing the monophony decoded frequency-domain signal without the energy adjustment.

[0009] An audio decoder, including: a judging unit, a processing unit, and a first reconstruction unit.

[0010] The judging unit is configured to judge whether bitstreams to be decoded are monophony coding layer and first stereo enhancement layer bitstreams. If the bitstreams to be decoded are the monophony coding layer and first

stereo enhancement layer bitstreams, the first reconstruction unit is triggered.

[0011] The processing unit is configured to decode the monophony coding layer to obtain a monophony decoded frequency-domain signal.

[0012] The first reconstruction unit is configured to reconstruct left and right channel frequency-domain signals in a first sub-band region by utilizing the monophony decoded frequency-domain signal after an energy adjustment, and reconstruct left and right channel frequency-domain signals in a second sub-band region by utilizing the monophony decoded frequency-domain signal without the energy adjustment, where the monophony decoded frequency-domain signal without the energy adjustment is obtained by the processing unit through decoding.

[0013] According to the embodiments of the present invention, a type of a monophonic signal used when the monophonic signal is reconstructed in a decoding process is determined according to a status of the bitstreams to be decoded. When it is determined that the bitstreams to be decoded are monophony coding layer and first stereo enhancement layer bitstreams, a monophony decoded frequency-domain signal after an energy adjustment is used to reconstruct left and right channel frequency-domain signals in a first sub-band region, and the monophony decoded frequency-domain signal without the energy adjustment is used to reconstruct left and right channel frequency-domain signals in a second sub-band region. The bitstreams to be decoded include only the monophony coding layer and first stereo enhancement layer bitstreams, and do not include a parameter of a residual in the second sub-band region. Therefore, the monophony decoded frequency-domain signal without the energy adjustment is used to reconstruct the left and right channel frequency-domain signals in the second sub-band region. In this way, signals at the coding end and the decoding end keep consistent, and quality of the decoded stereo signal is improved.

BRIEF DESCRIPTION OF THE DRAWINGS

[0014]

FIG. 1 is a flow chart of a parametric stereo audio coding method;
FIG. 2 is a flow chart of an audio decoding method according to an embodiment of the present invention;
FIG. 3 is a flow chart of another audio decoding method according to an embodiment of the present invention;
FIG. 4 is a schematic structural diagram of an audio decoder 1 according to an embodiment of the present invention;
and
FIG. 5 is a schematic structural diagram of an audio decoder 2 according to an embodiment of the present invention.

DETAILED DESCRIPTION OF THE EMBODIMENTS

[0015] The inventor of the present invention finds that: Quality of a stereo signal reconstructed by using a conventional audio decoding method depends on two factors: quality of a reconstructed monophonic signal and accuracy of an extracted stereo parameter. The quality of the monophonic signal reconstructed at a decoding end plays a very important part in the quality of a reconstructed stereo signal that is ultimately output. Therefore, the quality of the monophonic signal reconstructed at the decoding end needs to be as high as possible, based on which a high-quality stereo signal can be reconstructed.

[0016] An embodiment of the present invention provides an audio decoding method, which enables processed signals at a coding end and a decoding end to be consistent, thus quality of a decoded stereo signal may be improved. Embodiments of the present invention also provide a corresponding audio decoder.

[0017] For persons skilled in the art to better understand and implement the embodiments of the present invention, the following describes operations performed at the coding end in parametric stereo coding in detail. FIG. 1 is a flow chart of a parametric stereo audio coding method. The specific steps are as follows:

S11: Extract a channel parameter ITD according to original left and right channel signals, perform a channel delay adjustment on the left and right channel signals according to the ITD parameter, and perform down-mixing on the adjusted left and right channel signals to obtain a monophonic signal (also called a mixed signal, that is, an M signal) and a side signal (S signal).

[0018] Frequency-domain signals of the M signal and S signal within the [0~7kHz] frequency band respectively are $M\{m(0), m(1), \dots, m(N-1)\}$ and $S\{s(0), s(1), \dots, s(N-1)\}$. Frequency-domain signals of left and right channels within the [0~7kHz] frequency band are obtained according to formula (1) as $L\{l(0), l(1), \dots, l(N-1)\}$ and $R\{r(0), r(1), \dots, r(N-1)\}$.

$$l(i) = m(i) + s(i)$$

$$r(i) = m(i) - s(i)$$

...(1)

[0019] S 12: Divide the frequency-domain signals of the left and right channels into 8 sub-bands, extract, according to the sub-bands, left and right channel parameters ILDs: $W[band][l]$, $W[band][r]$, and quantize and code the parameters to obtain the quantized channel parameters ILDs: $W_q[band][l]$, $W_q[band][r]$, where $band \in (0, 1, 2, 3, 4, 5, 6, 7)$, l indicates the left channel parameter ILD, and r indicates the right channel parameter ILD.

[0020] S 13: Code the M signal and perform local decoding to obtain a locally decoded frequency-domain signal $M_1\{m_1(0), m_1(1), \dots, m_1(N-1)\}$.

[0021] S 14: Divide the M_1 frequency-domain signal obtained in S 13 into 8 sub-bands same as those of the left and right channels, compute an energy compensation parameter $ecomp[band]$ of sub-bands 5, 6, and 7 according to formula (2), and quantize and code the energy compensation parameter to obtain the quantized energy compensation parameter $ecomp_q[band]$.

$$ecomp[band] = \begin{cases} 10 \lg \left(\frac{C[band][l][l]}{W_q[band][l] \times W_q[band][l] \times Unmofiyenergy[band]} \right), & W_q[band][l] > 1 \\ 10 \lg \left(\frac{C[band][r][r]}{W_q[band][r] \times W_q[band][r] \times Unmofiyenergy[band]} \right), & W_q[band][l] \leq 1 \end{cases}$$

....(2)

[0022] In formula (2), $C[band][l][l] = \sum_{i \in [start_{band}, end_{band}]} l(i) \times l(i)$,

$C[band][r][r] = \sum_{i \in [start_{band}, end_{band}]} l(i) \times l(i)$, and $Unmofiyenergy[band] = \sum_{i \in [start_{band}, end_{band}]} m_1(i) \times m_1(i)$ re-

spectively indicate original left channel energy, original right channel energy, and locally decoded monophony energy that are in a current sub-band, and $[start_{band}, end_{band}]$ indicates a start position and an end position of a current sub-band frequency point.

[0023] S 15: Perform a frequency spectrum peak value analysis on the locally decoded frequency-domain signal M_1 to obtain a frequency spectrum analysis result $MASK\{mask(0), mask(1), \dots, mask(N-1)\}$, where $mask(i) \in \{0, 1\}$. If a frequency spectrum signal m_1 of M_1 in a position i is a peak value, $mask(i)=1$; if the frequency spectrum signal m_1 of M_1 in the position i is not a peak value, $mask(i)=0$.

[0024] S 16: Select an optimum energy adjusting factor multiplier, perform an energy adjustment on the decoded frequency-domain signal M_1 according to formula (3) to obtain a frequency-domain signal $M_2\{m_2(0), m_2(1), \dots, m_2(1), \dots, m_2(N-1)\}$ after the energy adjustment, and quantize and code the energy adjusting factor multiplier.

$$m_2(i) = \begin{cases} m_1(i) \times multiplier, & mask(i) = 0 \\ m_1(i) & , \quad mask(i) = 1 \end{cases} \dots\dots\dots(3)$$

[0025] S 17: Compute left and right channel residual information $resleft\{eleft(0), eleft(1), \dots, eleft(N-1)\}$ and $resright\{eright(0), eright(1), \dots, eright(N-1)\}$ according to formula (4) by utilizing the frequency-domain signal M_2 after the energy adjustment, left and right channel frequency-domain signals L and R , and the quantized channel parameter ILD W_q of the left and right channels.

$$\begin{aligned}
e_{left}(i) &= l(i) - W_q[band][l] \times m_2(i) \\
e_{right}(i) &= r(i) - W_q[band][r] \times m_2(i), \quad i \in [start_{band}, end_{band}], band = 0, 1, 2, 3, \dots, 7 \dots (4)
\end{aligned}$$

5

[0026] S 18: Perform a Karhunen-Loeve (K-L) transform on the left and right channel residuals, quantize and code a transform kernel H, and perform hierarchical and multiple quantizing and coding on a residual primary component $EU\{eu(0), eu(1), \dots, eu(N-1)\}$ and a residual secondary component $ED\{ed(0), ed(1), \dots, ed(N-1)\}$ that are obtained after the transform.

10

[0027] S19: Perform, according to the importance, hierarchical bitstream encapsulation on various coding information extracted at the coding end, and transmit a coding bitstream.

[0028] The coding information about the M signal is the most important, which is encapsulated as a monophony coding layer first; the channel parameters ILD and ITD, energy adjusting factor, energy compensation parameter, K-L transform kernel, and a first quantizing and coding result of the residual primary component in sub-bands 0 to 4 are encapsulated as a first stereo enhancement layer; other information is also encapsulated hierarchically according to the importance.

15

[0029] A network environment for bitstream transmission is changing all the time. If network resources are insufficient, not all coding information can be received at the decoding end. For example, only monophony coding layer and first stereo enhancement layer bitstreams are received, and bitstreams of other layers are not received.

20

[0030] During the process of researching and implementing the prior art, the inventor of the present invention finds that: In the case that only the monophony coding layer and first stereo enhancement layer bitstreams are received at the decoding end, that is, bitstreams to be decoded only include the monophony coding layer and first stereo enhancement layer bitstreams, energy compensation performed at the decoding end in the prior art is based on a monophony decoded frequency-domain signal after the energy adjustment, while extracting energy compensation parameters of sub-bands 5, 6, and 7 at the coding end in S 14 is based on a monophony decoded frequency-domain signal without the energy adjustment. Therefore, the processed signal at the coding end and the processed signal at the decoding end are inconsistent, and the inconsistency of the signals at the coding end and the decoding end cause quality of signals output after decoding to decline.

25

[0031] However, according to the embodiment of the present, a type of the monophony decoded frequency-domain signal used in the decoding process is determined according to a status of the bitstreams to be decoded at the decoding end. If only the monophony coding layer and first stereo enhancement layer bitstreams are received at the decoding end, the monophony decoded frequency-domain signal without the energy adjustment is used to reconstruct stereo signals of sub-bands 5, 6, and 7, while the monophony decoded frequency-domain signal after the energy adjustment is used to reconstruct stereo signals of sub-bands 0 to 4.

30

[0032] FIG. 2 is a flow chart of an audio decoding method according to an embodiment of the present invention, and the method includes:

35

S21: Determine that bitstreams to be decoded are monophony coding layer and first stereo enhancement layer bitstreams;

40

S22: Decode the monophony coding layer bitstream to obtain a monophony decoded frequency-domain signal;

S23: Reconstruct left and right channel frequency-domain signals in a first sub-band region by utilizing the monophony decoded frequency-domain signal after an energy adjustment; and

S24: Reconstruct left and right channel frequency-domain signals in a second sub-band region by utilizing the monophony decoded frequency-domain signal without the energy adjustment.

45

[0033] In the audio decoding method provided in the embodiment of the present invention, a type of a monophonic signal used when the monophonic signal is reconstructed in the decoding process is determined according to a status of the received bitstreams. After it is determined that the received bitstreams are the monophony coding layer and first stereo enhancement layer bitstreams, the monophony decoded frequency-domain signal after the energy adjustment is used to reconstruct left and right channel frequency-domain signals in a first sub-band region, and the monophony decoded frequency-domain signal without the energy adjustment is used to reconstruct left and right channel frequency-domain signals in a second sub-band region. The bitstreams to be decoded include only the monophony coding layer and first stereo enhancement layer bitstreams, and no parameter of a residual in the second sub-band region is received at a decoding end, so the monophony decoded frequency-domain signal without the energy adjustment is used to reconstruct the left and right channel frequency-domain signals in the second sub-band region. In this way, the processed signals at a coding end and the decoding end keep consistent, and therefore, quality of a decoded stereo signal may be improved.

50

55

[0034] FIG. 3 is a flow chart of another audio decoding method according to another embodiment of the present

invention. Through specific steps, the following describes in detail the decoding method used at the decoding end according to the embodiment of the present invention in a case that only monophony coding layer and first stereo enhancement layer bitstreams are received at the decoding end.

[0035] S31: Judge whether received bitstreams only include monophony coding layer and first stereo enhancement layer bitstreams. If the received bitstreams only include monophony coding layer and first stereo enhancement layer bitstreams, step S23 is executed.

[0036] S32: Use any audio/voice decoder corresponding to an audio/voice coder used at a coding end to decode the received monophony coding layer bitstream to obtain a monophony decoded frequency-domain signal: $M_1\{m_1(0), m_1(1), \dots, m_1(N-1)\}$, which is the signal obtained in S13 at the coding end, read a code word corresponding to each parameter from the first stereo enhancement layer bitstream, and decode each parameter to obtain channel parameters ILDs: $W_q[band][l]$, $W_q[band][r]$, a channel parameter ITD, an energy adjusting factor multiplier, a quantized energy compensation parameter $ecomp_q[band]$, a K-L transform kernel H, and a first quantizing result of a residual primary component in sub-bands 0 to 4 $EU_{q1}\{eu_{q1}(0), eu_{q1}(1), \dots, eu_{q1}(end_4), 0, 0, \dots, 0\}$.

[0037] S33: Perform a frequency spectrum peak value analysis on the monophony decoded frequency-domain signal M_1 , that is, search for a frequency spectrum maximum value in the frequency domain to obtain a frequency spectrum analysis result: $MASK\{mask(0), mask(1), \dots, mask(N-1)\}$, where $mask(i) \in \{0, 1\}$. If a frequency spectrum signal $m_1(i)$ of M_1 in a position i is a peak value, that is, the maximum value, $mask(i)=1$; if the frequency spectrum signal $m_1(i)$ of M_1 in a position i is not a peak value, $mask(i)=0$.

[0038] S34: Perform an energy adjustment on the monophony decoded frequency-domain signal by utilizing formula (5) according to the energy adjusting factor multiplier obtained through decoding and the frequency spectrum analysis result.

$$m_2(i) = \begin{cases} m_1(i) \times multiplier, & mask(i) = 0 \\ m_1(i), & mask(i) = 1 \end{cases} \dots\dots\dots(5)$$

[0039] In this way, the monophony decoded frequency-domain signal $M_2\{m_2(0), m_2(1), \dots, m_2(N-1)\}$ after the energy adjustment is obtained.

[0040] S35: Perform an anti-K-L transform according to formula (6) by utilizing the K-L transform kernel H and the first quantizing result of the residual primary component in the sub-bands 0 to 4 $EU_{q1}\{eu_{q1}(0), eu_{q1}(1), \dots, eu_{q1}(end_4), 0, 0, \dots, 0\}$, to obtain first quantizing residual information of the left and right channels in the sub-bands 0 to 4, that is, $resleft_{q1}\{eleft_{q1}(0), eleft_{q1}(1), \dots, eleft_{q1}(end_4), 0, 0, \dots, 0\}$ and $resright_{q1}\{eright_{q1}(0), eright_{q1}(1), \dots, eright_{q1}(end_4), 0, 0, \dots, 0\}$

$$\begin{bmatrix} resleft_{q1} \\ resright_{q1} \end{bmatrix} = H^{-1} \begin{bmatrix} eu_{q1} \\ 0 \end{bmatrix} \dots\dots\dots(6)$$

[0041] S36: Reconstruct left and right channel frequency-domain signals in the sub-bands 0 to 4 according to formula (7) by utilizing a monophony decoded frequency-domain signal M_2 after the energy adjustment, and reconstruct left and right channel frequency-domain signals in sub-bands 5, 6, and 7 according to formula (8) by utilizing the monophony decoded frequency-domain signal M_1 without the energy adjustment.

$$\begin{aligned} l'(i) &= eleft_{q1}(i) + W_q[band][l] \times m_2(i) \\ r'(i) &= eright_{q1}(i) + W_q[band][r] \times m_2(i), \quad i \in [start_{band}, end_{band}], band = 0, 1, 2, 3, 4 \dots\dots(7) \end{aligned}$$

$$\begin{aligned} l'(i) &= eleft_{q1}(i) + W_q[band][l] \times m_1(i) \\ r'(i) &= eright_{q1}(i) + W_q[band][r] \times m_1(i), \quad i \in [start_{band}, end_{band}], band = 5, 6, 7 \dots\dots\dots(8) \end{aligned}$$

[0042] The first stereo enhancement layer bitstream that includes the left and right channel residual information in the sub-bands 0 to 4 is received at the decoding end, so the monophony decoded frequency-domain signal M2 after the energy adjustment is used to reconstruct the left and right channel frequency-domain signals when stereo signals of sub-bands 0 to 4 are reconstructed. The decoding end does not receive any other enhancement layer bitstreams except the monophony coding layer and first stereo enhancement layer bitstreams, so that left and right channel residual information in the sub-bands 5, 6, and 7 cannot be obtained. Moreover, in S 14 at the coding end, the energy compensation parameters of the sub-bands 5, 6, and 7 are extracted according to formula (2), and it may be seen from S14 that, the energy compensation parameters are based on the monophony decoded frequency-domain signal M1, so that the monophony decoded frequency-domain signal M₁ without the energy adjustment is used for reconstruction when the stereo signals of the sub-bands 5, 6, and 7 are reconstructed in this step, while the monophony decoded frequency-domain signal M₂ after the energy adjustment is used for reconstruction when the stereo signals of the sub-bands 0 to 4 are reconstructed, thus signals at the coding end and decoding end keep consistent.

[0043] S37: Perform an energy compensation adjustment on the sub-bands 5, 6, and 7 of the reconstructed left and right channel frequency-domain signals according to formula (9).

$$l'(i) = l'(i) \times 10^{ecomp_q[band]/20}$$

$$r'(i) = r'(i) \times 10^{ecomp_q[band]/20}, \quad i \in [start_{band}, end_{band}], band = 5, 6, 7 \dots \dots \dots (9)$$

[0044] S38: Process the left and right channel frequency-domain signals to obtain the ultimate left and right channel output signals.

[0045] In the preceding parametric stereo audio coding process, frequency-domain signals are divided into 8 sub-bands, sub-bands 0 to 4 of primary component parameters are encapsulated at the first stereo enhancement layer, and other parameters related to the residual are encapsulated at other stereo enhancement layers. It should be noted that the sub-bands 0 to 4 are referred to as the first sub-band region, and the sub-bands 5 to 7 are referred to as the second sub-band region here. It may be understood that, in specific implementation, frequency-domain signals may also be divided into multiple, other than 8, sub-bands in a parametric stereo audio coding process. Even if frequency-domain signals are divided into 8 sub-bands, the 8 sub-bands may also be divided into two sub-band regions different from the foregoing. For example, the sub-bands 0 to 3 of primary component parameters are encapsulated at the first stereo enhancement layer, and other parameters related to the residual are encapsulated at other stereo enhancement layers, so that in this case, the sub-bands 0 to 3 are referred to as a first sub-band region, and the sub-bands 4 to 7 are referred to as a second sub-band region. Correspondingly, in the case that bitstreams to be decoded only include monophony coding layer and first stereo enhancement layer bitstreams, according to the embodiment of the present invention, the monophony decoded frequency-domain signal after the energy adjustment is used to reconstruct left and right channel frequency-domain signals in the sub-bands 0 to 3 (the first sub-band region) at the decoding end, and the monophony decoded frequency-domain signal without the energy adjustment is used to reconstruct the left and right channel frequency-domain signals in the sub-bands 4 to 7 (the second sub-band region).

[0046] It may be seen from the embodiment that, the type of the monophonic signal used when a monophonic signal is reconstructed in the decoding process is determined according to the status of the received bitstreams. When it is determined that the received bitstreams are the monophony coding layer and first stereo enhancement layer bitstreams, the monophony decoded frequency-domain signal after the energy adjustment is used to reconstruct the left and right channel frequency-domain signals in the first sub-band region, and the monophony decoded frequency-domain signal without the energy adjustment is used to reconstruct the left and right channel frequency-domain signals in the second sub-band region. The bitstreams to be decoded only include the monophony coding layer and first stereo enhancement layer bitstreams, and no parameter of the residual in the second sub-band region is received at the decoding end, so that the monophony decoded frequency-domain signal without the energy adjustment is used to reconstruct the left and right channel frequency-domain signals in the second sub-band region. In this way, the processed signals at the coding end and the decoding end keep consistent, and therefore, quality of a decoded stereo signal may be improved.

[0047] In the case that the decoding end also receives other stereo enhancement layer bitstreams (for example, all bitstreams of the monophony coding layer and all stereo enhancement layers are received) besides the monophony coding layer and first stereo enhancement layer bitstreams, the decoding process is different from the foregoing process. The difference lies in that residual information in all sub-band regions may be obtained through decoding. Therefore, the monophony decoded frequency-domain signal after the energy adjustment is used to reconstruct the left and right channel frequency-domain signals (including stereo signals in the first and second sub-band regions). In addition, the complete residual signals in all sub-band regions can be obtained, therefore, energy compensation does not need to be performed on the left and right channel frequency-domain signals in the first or second sub-band. In this way, processed

signals at the coding end and decoding end are consistent.

[0048] The audio decoding method according to the embodiment of the present invention is described above in detail. The following correspondingly describes a decoder that uses the foregoing audio decoding method.

[0049] FIG. 4 is a schematic structural diagram of an audio decoder 1 according to an embodiment of the present invention, and the audio decoder 1 includes: a judging unit 41, a processing unit 42, and a first reconstruction unit 43.

[0050] The judging unit 41 is configured to judge whether bitstreams to be decoded are a monophony coding layer and first stereo enhancement layer bitstreams. If the bitstreams to be decoded are the monophony coding layer and the first stereo enhancement layer bitstreams, the first reconstruction unit 43 is triggered.

[0051] The processing unit 42 is configured to decode the monophony coding layer to obtain a monophony decoded frequency-domain signal.

[0052] The first reconstruction unit 43 is configured to reconstruct left and right channel frequency-domain signals in a first sub-band region by utilizing the monophony decoded frequency-domain signal after an energy adjustment, and reconstruct left and right channel frequency-domain signals in a second sub-band region by utilizing the monophony decoded frequency-domain signal without the energy adjustment, where the monophony decoded frequency-domain signal without the energy adjustment is obtained by the processing unit 42 through decoding.

[0053] The processing unit 42 is further configured to decode the first stereo enhancement layer bitstream to obtain an energy adjusting factor, perform a frequency spectrum peak value analysis on the monophony decoded frequency-domain signal to obtain a frequency spectrum analysis result, and perform an energy adjustment on the monophony decoded frequency-domain signal according to the frequency spectrum analysis result and the energy adjusting factor.

[0054] If in a parametric stereo audio coding process, frequency-domain signals are divided into 8 sub-bands, sub-bands 0 to 4 of a primary component parameter are encapsulated at a first stereo enhancement layer, and other parameters related to a residual are encapsulated at other stereo enhancement layers, the first reconstruction unit 43 is specifically configured to use the monophony decode frequency-domain signal after the energy adjustment to reconstruct the left and right channel frequency-domain signals in sub-bands 0 to 4, and use the monophony decode frequency-domain signal without the energy adjustment to reconstruct the left and right channel frequency-domain signals in sub-bands 5, 6, and 7, where the monophony decode frequency-domain signal without the energy adjustment is derived by the processing unit 42 through decoding.

[0055] After the first reconstruction unit 43 obtains the reconstructed left and right channel frequency-domain signals, the processing unit 42 is further configured to perform an energy compensation adjustment on sub-bands 5, 6, and 7 of the reconstructed left and right channel frequency-domain signals.

[0056] It can be seen that, after determining that only a monophony coding layer and first stereo enhancement layer bitstreams are received, the audio decoder introduced in this embodiment uses the monophony decoded frequency-domain signal after the energy adjustment to reconstruct the left and right channel frequency-domain signals in the first sub-band region, and uses the monophony decoded frequency-domain signal without the energy adjustment to reconstruct the left and right channel frequency-domain signals in a second sub-band region. Only the monophony coding layer and first stereo enhancement layer bitstreams are received, so that no parameter of the residual in the second sub-band region is received. Therefore, the monophony decoded frequency-domain signal without the energy adjustment is used to reconstruct the left and right channel frequency-domain signals in the second sub-band region. In this way, processed signals at the decoding end and the coding end keep consistent, and therefore, quality of a decoded stereo signal may be improved.

[0057] FIG. 4 is a schematic structural diagram of an audio decoder 2 according to an embodiment of the present invention. Different from the audio decoder 1, the audio decoder 2 further includes a second reconstruction unit 51.

[0058] When a judging result of the judging unit 41 is that in addition to a monophony coding layer and first stereo enhancement layer bitstreams, bitstreams to be decoded further include other stereo enhancement layer bitstreams, the second reconstruction unit 51 is configured to use the monophony decode frequency-domain signal after the energy adjustment to reconstruct left and right channel frequency-domain signals in all sub-band regions.

[0059] It may be understood that, in specific implementation, the first reconstruction unit 43 and the second reconstruction unit 51 may be integrated to be used as one reconstruction unit.

[0060] Persons of ordinary skill in the art may understand that all or part of the steps of the method according to the foregoing embodiments may be implemented by a program instructing relevant hardware. The program may be stored in a computer readable storage medium. The storage medium may be a Read-Only Memory (ROM), a Random Access Memory (RAM), a magnetic disk or an optical disk.

[0061] The audio processing method and the audio decoder provided in the embodiments of the present invention are described in detail above. The principle and implementation of the present invention are described through specific examples. The description about the foregoing embodiments is merely used to help understand the method and core ideas of the present invention. Meanwhile, persons of ordinary skill in the art may make variations and modifications to the present invention in terms of the specific implementations and application scopes according to the ideas of the present invention. Therefore, the specification shall not be construed as limitations to the present invention.

Claims**1.** An audio decoding method, comprising:

5 determining that bitstreams to be decoded are monophony coding layer and first stereo enhancement layer bitstreams;
 decoding the monophony coding layer bitstream to obtain a monophony decoded frequency-domain signal;
 reconstructing left and right channel frequency-domain signals in a first sub-band region by utilizing the monophony decoded frequency-domain signal after an energy adjustment; and
 10 reconstructing left and right channel frequency-domain signals in a second sub-band region by utilizing the monophony decoded frequency-domain signal without the energy adjustment.

2. The method according to claim 1, further comprising:

15 performing the energy adjustment on the monophony decoded frequency-domain signal.

3. The method according to claim 2, wherein the performing the energy adjustment on the monophony decoded frequency-domain signal comprises:

20 decoding the first stereo enhancement layer bitstream to obtain an energy adjusting factor;
 performing a frequency spectrum peak value analysis on the monophony decoded frequency-domain signal to obtain a frequency spectrum analysis result; and
 performing the energy adjustment on the monophony decoded frequency-domain signal according to the frequency spectrum analysis result and the energy adjusting factor.

4. The method according to any one of claims 1 to 3, wherein the reconstructing the left and right channel frequency-domain signals by utilizing the monophony decoded frequency-domain signal after the energy adjustment in the first sub-band region; and the reconstructing the left and right channel frequency-domain signals by utilizing the monophony decoded frequency-domain signal without the energy adjustment in the second sub-band region specifically comprise:

25 using the monophony decoded frequency-domain signal after the energy adjustment to reconstruct the left and right channel frequency-domain signals in sub-bands 0 to 4, and using the monophony decoded frequency-domain signal without the energy adjustment to reconstruct the left and right channel frequency-domain signals in sub-bands 5, 6, and 7.

5. The method according to claim 4, wherein after the reconstructing the left and right channel frequency-domain signals, the method further comprises:

35 performing an energy compensation adjustment on the sub-bands 5, 6, and 7 of the reconstructed left and right channel frequency-domain signals.

6. An audio decoder, comprising a judging unit, a processing unit, and a first reconstruction unit, wherein:

45 the judging unit is configured to judge whether bitstreams to be decoded are monophony coding layer and first stereo enhancement layer bitstreams, and if the bitstreams to be decoded are the monophony coding layer and first stereo enhancement layer bitstreams, the first reconstruction unit is triggered;
 the processing unit is configured to decode the monophony coding layer to obtain a monophony decoded frequency-domain signal; and
 50 the first reconstruction unit is configured to reconstruct left and right channel frequency-domain signals in a first sub-band region by utilizing the monophony decoded frequency-domain signal after an energy adjustment, and reconstruct the left and right channel frequency-domain signals in a second sub-band region by utilizing the monophony decoded frequency-domain signal without the energy adjustment, wherein the monophony decoded frequency-domain signal without the energy adjustment is obtained by the processing unit through decoding.

7. The audio decoder according to claim 6, wherein the processing unit is further configured to decode the first stereo enhancement layer bitstream to obtain an energy adjusting factor, perform a frequency spectrum peak value analysis on the monophony decoded frequency-domain signal to obtain a frequency spectrum analysis result, and perform

the energy adjustment on the monophony decoded frequency-domain signal according to the frequency spectrum analysis result and the energy adjusting factor.

5 **8.** The audio decoder according to claim 7, wherein the first reconstruction unit is specifically configured to reconstruct the left and right channel frequency-domain signals in sub-bands 0 to 4 by utilizing the monophony decoded frequency-domain signal after the energy adjustment, and reconstruct the left and right channel frequency-domain signals in sub-bands 5, 6, and 7 by utilizing the monophony decoded frequency-domain signal without the energy adjustment, wherein the monophony decoded frequency-domain signal without the energy adjustment is obtained by the processing unit through decoding.

10 **9.** The audio decoder according to claim 8, wherein after the first reconstruction unit obtains the reconstructed left and right channel frequency-domain signals, the processing unit is further configured to perform an energy compensation adjustment on the sub-bands 5, 6, and 7 of the reconstructed left and right channel frequency-domain signals.

15 **10.** The audio decoder according to claim 6, further comprising a second reconstruction unit, wherein when a judging result of the judging unit is that in addition to the monophony coding layer and first stereo enhancement layer bitstreams, the bitstreams to be decoded further comprise other stereo enhancement layer bitstreams, and the second reconstruction unit is configured to use the monophony decoded frequency-domain signal after the energy adjustment to reconstruct left and right channel frequency-domain signals in all sub-band regions.

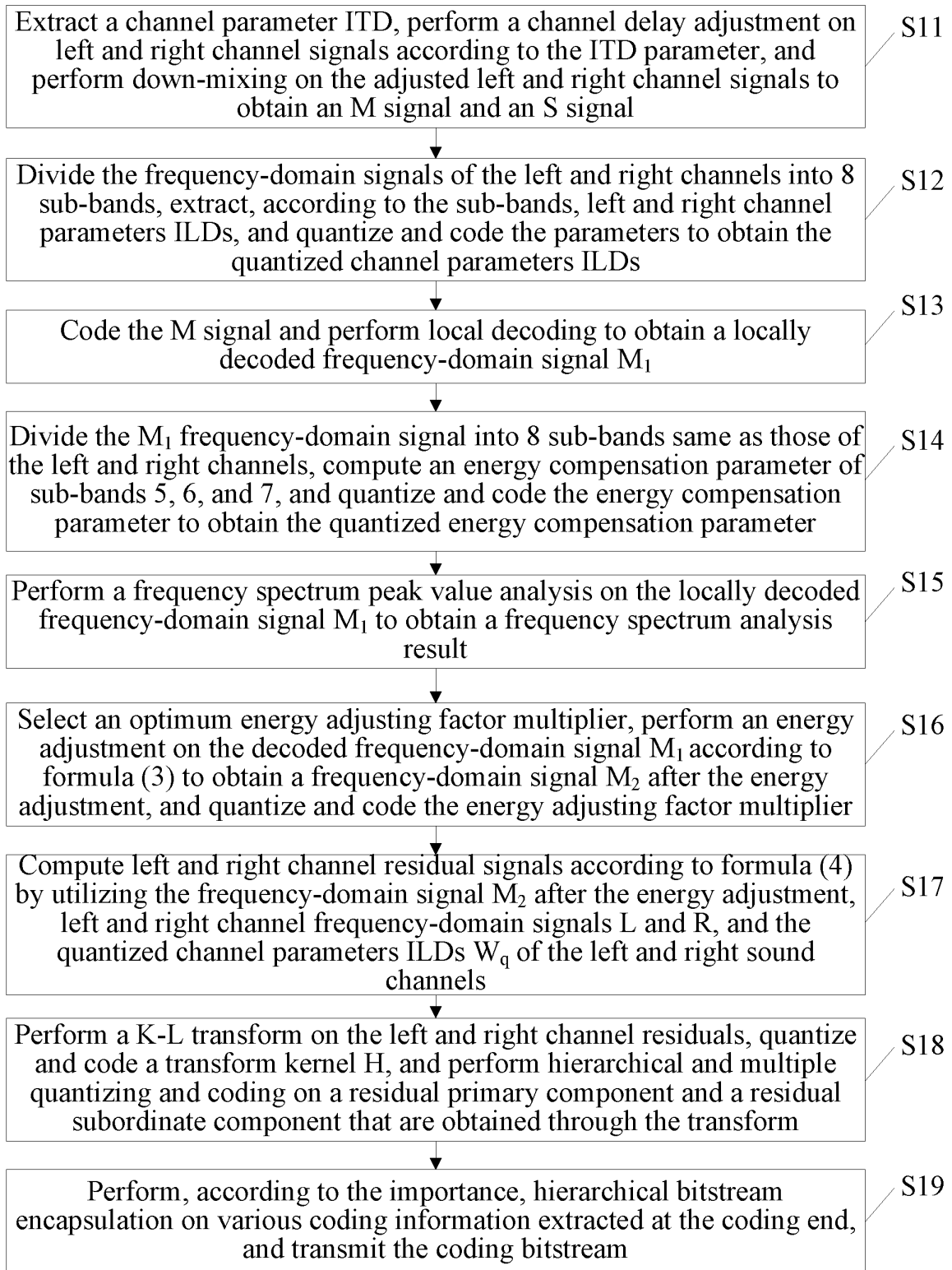


FIG. 1

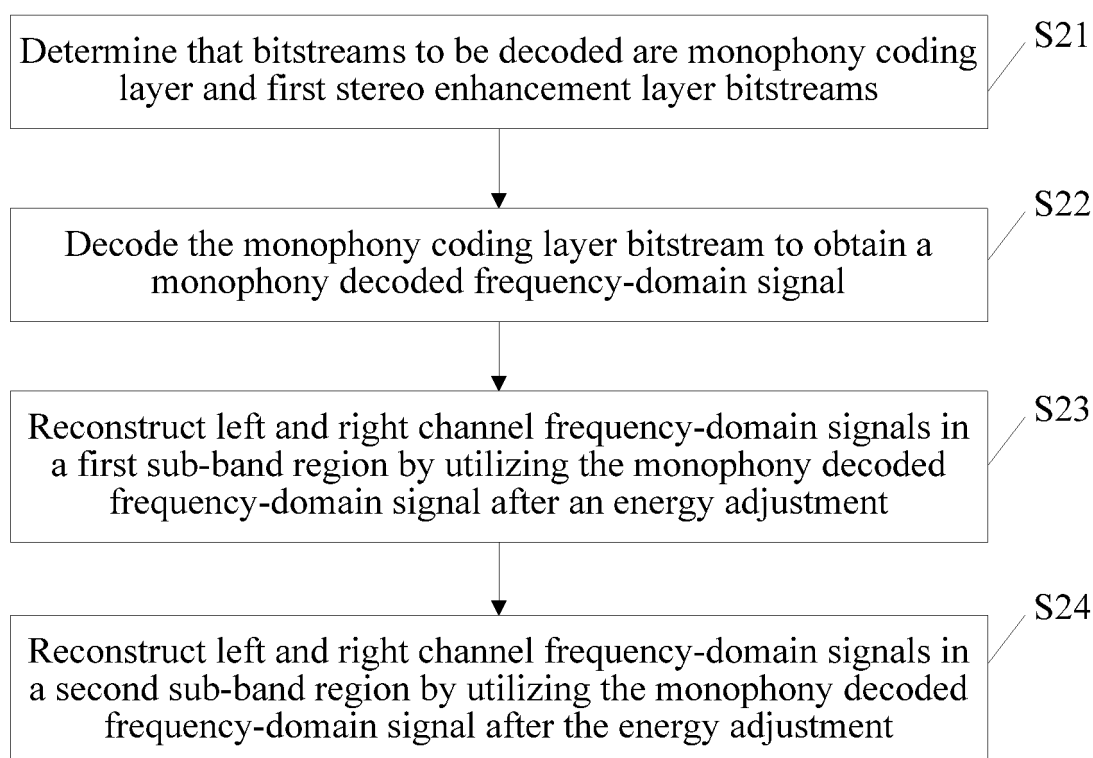


FIG. 2

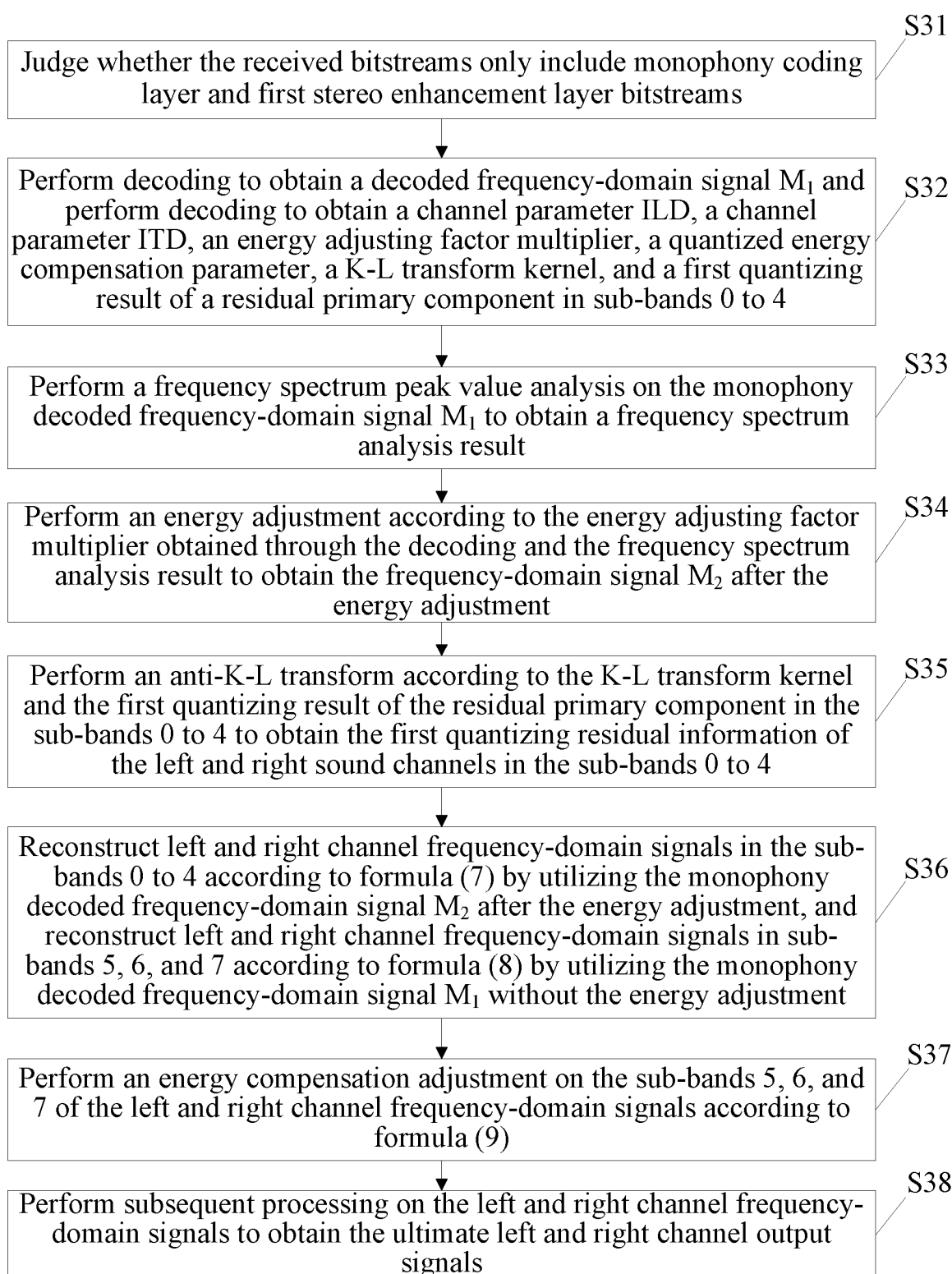


FIG. 3

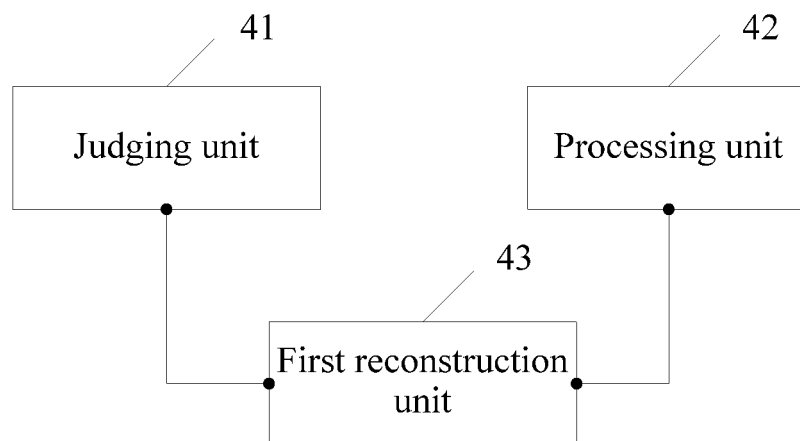


FIG. 4

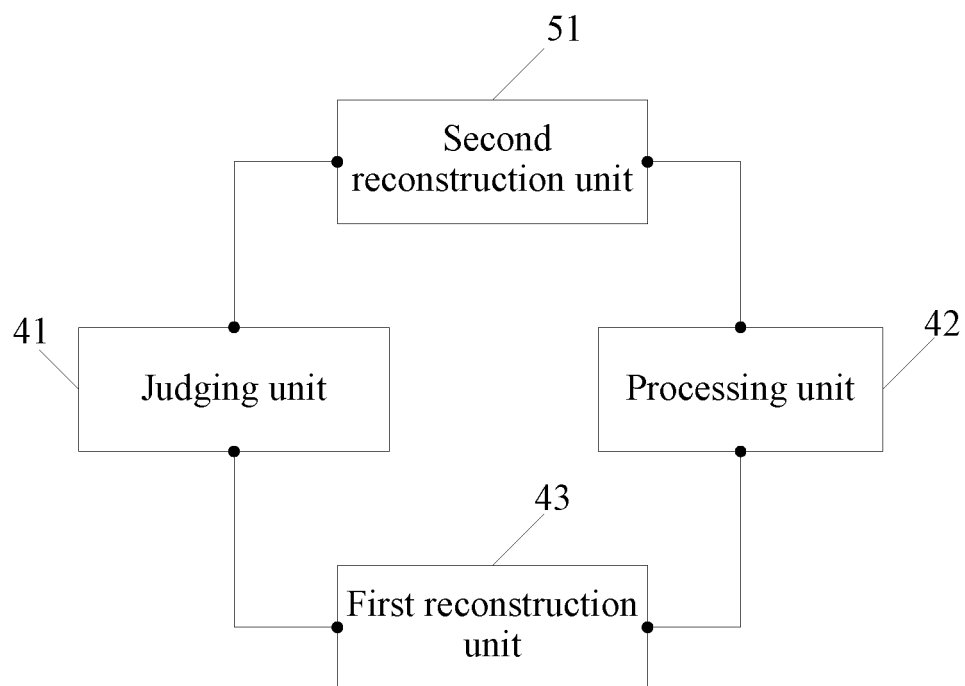


FIG. 5

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2010/072781

A. CLASSIFICATION OF SUBJECT MATTER

See extra sheet

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC: G10L, H04S, H04H, G10K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNPAT, CNKI, WPI, EPODOC, PAJ: sound, voice, speech, audio, decod+,
monophony, track, stereo, reconstruct+, restor+, reproduce+**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN1875402A(KONINKL. PHILIPS ELECTRONICS NV), 06 Dec. 2006(06.12.2006), page 1, paragraph 2, figure 1	1-10
A	CN101433099A(ERICSSON TELEFON AB L M), 13 May 2009(13.05.2009), the whole document	1-10
A	WO02/091362A1(FRANCE TELECOM), 14 Nov. 2002(14.11.2002), the whole document	1-10
A	JP11-18199A(KAWAI MUSICAL INSTR MFG CO), 22 Jan. 1999(22.01.1999), the whole document	1-10
A	US2008/0161952A1(KABUSHIKI KAISHA TOSHIBA), 03 July 2008(03.07.2008), the whole document	1-10

☒ Further documents are listed in the continuation of Box C.☒ See patent family annex.

* Special categories of cited documents:	“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
“A” document defining the general state of the art which is not considered to be of particular relevance	“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
“E” earlier application or patent but published on or after the international filing date	“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
“L” document which may throw doubts on priority claim (S) or which is cited to establish the publication date of another citation or other special reason (as specified)	“&” document member of the same patent family
“O” document referring to an oral disclosure, use, exhibition or other means	
“P” document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search

11. Aug. 2010(11.08.2010)

Date of mailing of the international search report

02 Sep. 2010 (02.09.2010)Name and mailing address of the ISA/CN
The State Intellectual Property Office, the P.R.China
6 Xitucheng Rd., Jimen Bridge, Haidian District, Beijing, China
100088
Facsimile No. 86-10-62019451

Authorized officer

ZHANG Yaling

Telephone No. (86-10)62085722

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2010/072781

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US6032081A(KAREA TELECOMMUNICATION AUTHORITY),29 Feb. 2000(29.02.2000), the whole document	1-10
A	CN101366321A(NOKIA CORP),11 Feb. 2009(11.02.2009), the whole document	1-10
A	ERIKSCHUIJERS, WERNER OOMEN, BERT DEN BRINKER, JEROEN BREEBAART: "Advances in Parametric Coding for High-Quality Audio" AUDIO ENGINEERING SOCIETY 114th, CONVENTION PAPER 5852, AMSTERDAM, THE NETHERLANDS, March 22-25 2003, pages 1-10	1-10

Form PCT/ISA /210 (continuation of second sheet) (July 2009)

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2010/072781

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN1875402A	06.12.2006	WO2005043511A1	12.05.2005
		US2007067162A1	22.03.2007
		US2007067162A1	22.03.2007
		US2009216544A1	27.08.2009
		JP2009163257A	23.07.2009
		EP1683133A1	26.07.2006
		DE602004004818T2	06.12.2007
		JP2007514962T	07.06.2007
		AT354160T	15.03.2007
		BRPI0415951A	02.01.2007
		CN101540171A	23.09.2009
		ES2282899T3	16.10.2007
		KR20060131745A	20.12.2006
		RU2374703C2	27.11.2009
		RU2374703C2	27.11.2009
CN101433099A	13.05.2009	WO2007078254A2	12.07.2007
		EP1969901A2	17.09.2008
		RU2008132156A	10.02.2010
WO02/091362 A1	14.11.2002	FR2824432A1	08.11.2002
JP11-18199A	22.01.1999	JP4029076B	15.05.1992
		JP1934633C	26.05.1995
US2008/0161952A1	03.07.2008	JP2008164823A	17.07.2008
US6032081A	29.02.2000	JP9167154A	24.06.1997
		KR0174084B1	01.04.1999
CN101366321A	11.02.2009	US2007160218A1	12.07.2007
		WO2007080211A1	19.07.2007
		US2007160219A1	12.07.2007
		EP1971979A1	24.09.2008
		EP1972180A1	24.09.2008
		JP2009522895T	11.06.2009

Form PCT/ISA /210 (patent family annex) (July 2009)

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2010/072781

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
		JP2009522894T	11.06.2009
		AU2007204332A1	19.07.2007
		AU2007204333A1	19.07.2007
		CA2635024A1	19.07.2007
		CA2635985A1	19.07.2007
		CN101366081A	11.02.2009
		KR20080074223A	12.08.2008
		KR20080078882A	12.08.2008
		RU2008126699A	20.02.2010
		RU2008127062A	20.02.2010

Form PCT/ISA /210 (patent family annex) (July 2009)

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2010/072781

Classification of subject matter

G10L 19/02(2006.01)i

H04S 3/00(2006.01)i

H04H 20/88(2008.01)i

REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Patent documents cited in the description

- CN 200910137565 [0001]